

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 14:23:37 ; Search time 3958 Seconds
(without alignments)
16051.706 Million cell updates/sec

Title: US-09-745-506-74

Perfect score: 1553

Sequence: 1 GTGATTGTATCTGTGTCCT.....TCCTGTACTTAACATTCMA 1553

Scoring table: OLIGO.NUC

Gapop 60.0 , Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBdb1:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_fod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1324	85.3	1574	6	BD158389
2	1324	85.3	1574	6	AK023378
3	1324	85.3	1606	9	AB038949
4	1138	73.3	1425	9	AF283538
5	1138	73.3	1440	9	BC007654
6	1137	73.2	1387	9	AF060513
7	1053	67.8	1053	6	AX119075
8	1027	66.1	1353	9	HSMB05522
9	852	54.9	1579	9	AF182416
10	567	36.5	796	6	BD149184
11	414	26.7	170586	2	AC037455
12	414	26.7	190508	9	AC005037
13	414	26.7	198250	2	AC093681
14	210	13.5	249	6	BD049005
15	141	9.1	198250	2	AC093681
16	90	5.8	339	6	BD040994
17	50	3.2	55374	6	AL645474
18	41	2.6	131	6	BD037535
19	38	2.4	1836	10	AF284439
20	38	2.4	149819	2	AC121091
21	38	2.4	231600	2	AC130779
22	38	2.4	234976	2	AC118698
23	22	1.4	11976	6	AX345497
24	22	1.4	155127	2	AL929495
25	21	1.4	66761	2	AC044823
26	21	1.4	77223	2	AC036228
27	21	1.4	96481	9	AC104811
28	21	1.4	102223	9	AC022432
29	21	1.4	104511	9	AL161638
30	21	1.4	134963	9	HS199116
31	21	1.4	142629	2	AP002479
32	21	1.4	151976	9	AC024568
33	21	1.4	153276	9	AC012569
34	21	1.4	157073	2	AC102733
35	21	1.4	157836	2	AC060228
36	21	1.4	160815	2	AP001895
37	21	1.4	163192	2	AC023412
38	21	1.4	163372	2	AC087744
39	21	1.4	163865	2	AC102221
40	21	1.4	164179	2	AL445533
41	21	1.4	179498	2	AC104921
42	21	1.4	180015	10	AC121925
43	21	1.4	181869	2	BX511267
44	21	1.4	185969	2	AC069037
45	21	1.4	187080	2	AC010308

ALIGNMENTS

RESULT 1
BD158389 1574 bp DNA linear PAT 17-JAN-2003
LOCUS
BD158389
DEFINITION
BD158389
ACCESSION
BD158389
VERSION
BD158389.1 GI:27864147
KEYWORDS
JP 2002191363-A/13232.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1574)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
AUTHORS
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE
Primer for synthesizing full-length cDNA and use thereof

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: JP 2002191363-A 13232 09-JUL-2002;
 COMMENT HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/13232
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO,
 PI JUNICHI YANAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C1201/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 FT CDS (221)..(1320).
 1..1574
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 420 a 361 c 372 g 421 t
 ORIGIN
 Query Match 85.3%; Score 1324; DB 6; Length 1574;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGATTGTATCTTGGTGTCTGACAGAGACAGACAGAGAGAGATTGGGTCAAGAAATCTC 60
 DB 27 GTGATTGTATCTTGGTGTCTGACAGAGACAGACAGAGAGAGATTGGGTCAAGAAATCTC 86
 QY 61 CCTCCGACACAGACAGACAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 DB 87 CCTCCGACACAGACAGACAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 146
 QY 121 CTGTGTCTCGT 180
 DB 147 CTGTGTCTCGT 206
 QY 181 TAGAT 240
 DB 207 TAGAT 266
 QY 241 CTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 DB 267 CTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 326
 QY 301 TGAGAGTTGGAGCAATGTTGGATTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
 DB 327 TGAGAGTTGGAGCAATGTTGGATTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 386
 QY 361 ACTTCTCTGACGACATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 420
 DB 387 ACTTCTCTGACGACATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 446
 QY 421 AGACCTATCTCTCTACACATCCGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 DB 447 AGACCTATCTCTCTACACATCCGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 506
 QY 481 CACATGAG 540
 DB 507 CACATGAG 566
 QY 541 TCATACAG 600
 DB 567 TCATACAG 626
 QY 601 AGCTTGTACCTCCAG 660
 DB 627 AGCTTGTACCTCCAG 686

QY 661 CCACCGAGTAGAATTCACGTTAACTACACCAAGACCTGGACAAAGTATGTGTGACT 720
 DB 687 CCACCGAGTAGAATTCACGTTAACTACACCAAGACCTGGACAAAGTATGTGTGACT 746
 QY 721 GAAGGAAATGAGCGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 780
 DB 747 GAAGGAAATGAGCGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 806
 QY 781 AACACGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 840
 DB 807 AACACGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 866
 QY 841 CCGAACAACAACTTTATTCAGAGAGAGAAATTCGTCTGCTGAGAGAAAGCTTGTCTCT 900
 DB 867 CCGAACAACAACTTTATTCAGAGAGAGAAATTCGTCTGCTGAGAGAAAGCTTGTCTCTCT 926
 QY 901 ACATACGTAATGAG 960
 DB 927 ACATACGTAATGAG 986
 QY 961 TGATCGAATTAATAAAGACACCTAAACATATCTCATATTCGTTAGCCCTGGGGTGGGAG 1020
 DB 987 TGATCGAATTAATAAAGACACCTAAACATATCTCATATTCGTTAGCCCTGGGGTGGGAG 1046
 QY 1021 AACCTTAGAGTCTCAAGTCAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 DB 1047 AACCTTAGAGTCTCAAGTCAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
 QY 1081 GCAGGGGTGAG 1140
 DB 1107 GCAGGGGTGAG 1166
 QY 1141 TGCTGTCTCCCAAGAGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1200
 DB 1167 TGCTGTCTCCCAAGAGATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1226
 QY 1201 TCTTTCTGACCTTGTGAGATATGCTGAGATCTTCATCTTGGAGATTAATGTAATGTAATGTAATGTAAT 1260
 DB 1227 TCTTTCTGACCTTGTGAGATATGCTGAGATCTTCATCTTGGAGATTAATGTAATGTAATGTAATGTAAT 1286
 QY 1261 ATCAGAGACTGACAG 1320
 DB 1287 ATCAGAGACTGACAG 1346
 QY 1321 ATTC 1324
 DB 1347 ATTC 1350
 RESULT 2
 AK023378
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ13316 f1s, clone OVAR1001555, weakly similar to NG1-INTERACTING FACTOR 3.
 ACCESSION AK023378
 VERSION AK023378.1 GI:10435289
 KEYWORDS Oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Magatsuma,M., Hosofri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahashi,K., Masuo,Y., Minomiya,K. and Iwayanagi,T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 1574)
 AUTHORS Isogai,T. and Otsuki,T.

TITLE

Martindale, D., Koop, B.F., Scherer, S.W., Nicholson, D.W.,
Rouleau, G.A., Ikeda, J.-E. and Hayden, M.R.
Cloning and characterization of three novel genes, ALS2CR1,
ALS2CR2, and ALS2CR3, in the juvenile amyotrophic lateral sclerosis
(ALS2) critical region at chromosome 2q33-q34: candidate genes for
ALS2

JOURNAL
Genomics 71 (2), 200-213 (2001)

MEDLINE
21100893

PUBMED
11161814

REFERENCE
2 (bases 1 to 1606)

AUTHORS
Hadano, S., Ikeda, J. and Hayden, M.R.

JOURNAL
Direct Submission

Submitted (24-FEB-2000) Shinji Hadano, Tokai University, The
Institute of Medical Sciences, Bohseidai, Isehara, Kanagawa
259-1193, Japan (E-mail: shinji.enga.med.u-tokai.ac.jp,
Tel.: +81-463-91-5095, Fax: +81-463-91-4993)

FEATURES
Location/Qualifiers

1..1606

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="2"

/map="2q33"

/note="IMAGE:32584-uniGene ID: Hs.21943"

1..1606

/gene="ALS2CR1"

286..1338

/gene="ALS2CR1"

/note="candidate 1"

/codon_start=1

/product="amyotrophic lateral sclerosis 2"

/protein_id="BAB32499.1"

/db_xref="GI:12862478"

/translation="MDKLKSLNDLFAISLFAESMDVNGLLVEPSPHVTNLEFLN
DLTEVMEVLOKRAADLISYHPLEPRMKTMTWTKERLIVIRALEKRGVSPHTA
YDAAPQGVNMLARGIGACTSRPIHPSAPRPYPRGNHREVENNYDGLKYSAYK
GIDSVSYSEFARNGNEBQTRINLNCOTKALMOYDPLSRKOLKQETIISLEPPL
LHTGMRICLIDESVSLATMDRIKRLKLSHIRIALGVRLSEQVAVVALCAGSGS
SLQGVENDLVLTEGMSHHDPLDAASQGINVILCHSNTBERGLSDRLMDLSLENK
INILSETDRDPLVYV"

polyA_signal
1566..1571

/gene="ALS2CR1"

BASE COUNT 438 a 365 c 375 g 428 t

ORIGIN

Query Match 85.3%; Score 1324; DB 9; Length 1606;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 GTGATGTTATCTGTGCTGCTGACAGAGACAGAGAGAGAGATTGGCTCAGAAAACCTGC 60
    |||||
DB 42 GTGATGTTATCTGTGCTGCTGACAGAGACAGAGAGAGAGATTGGCTCAGAAAACCTGC 101
OY 61 CCTGCCGACACAGACAGACAGCAGCTAGTGGAGACAGGGGCTCTGACTCAGACTTAACG 120
    |||||
DB 102 CCTGCCGACACAGACAGACAGCAGCTAGTGGAGACAGGGGCTCTGACTCAGACTTAACG 161
OY 121 CTGTCTCTGTGTTTTCACCTGCTCTGAAAAAGGCTGAAGTGCACATGAATGAGCA 180
    |||||
DB 162 CTGTCTCTGTGTTTTCACCTGCTCTGAAAAAGGCTGAAGTGCACATGAATGAGCA 221
OY 181 TAGATGATCCCGACAGACAGATCCGTTGTAGATTCCTGATCTGCAATCTTCCCGTT 240
    |||||
DB 222 TAGATGATCCCGACAGACAGATCCGTTGTAGATTCCTGATCTGCAATCTTCCCGTT 281
OY 241 CTTCATGATTTGAAGGCTCCTCTTCTTCTGATGATGATGATGATGATGATGATGATG 300
    |||||
DB 282 CTTCATGATTTGAAGGCTCCTCTTCTTCTGATGATGATGATGATGATGATGATGATG 341
OY 301 TGAGAGTTGGAGCAATGTTGATTTACTGTGTGAGAACCAAGCCACACATCTGTAATAC 360
    |||||
DB 342 TGAGAGTTGGAGCAATGTTGATTTACTGTGTGAGAACCAAGCCACACATCTGTAATAC 401

```

```

OY 361 ACTCTTCCTGACCAATGACTGACTGAGAGAGTGTGAGGAGGCTGCAAAAAGAGCC 420
    |||||
DB 402 ACTCTTCCTGACCAATGACTGACTGAGAGAGTGTGAGGAGGCTGCAAAAAGAGCC 461
OY 421 AGACCTATTTCTCTCTCCATCCGCTATCTTCCGACCCATGAAGCCATTAACCTGGA 480
    |||||
DB 462 AGACCTATTTCTCTCTCCATCCGCTATCTTCCGACCCATGAAGCCATTAACCTGGA 521
OY 481 CACATGAGAGAGAGCCCTGATGATCCGGGCTGTGAGAGAGAGAGCTATCTACTCTCC 540
    |||||
DB 522 CACATGAGAGAGAGCCCTGATGATCCGGGCTGTGAGAGAGAGAGCTATCTACTCTCC 581
OY 541 TCATACACCCATGATGATGCTGCGCCCAAGGGGCTCAACACTGTTGGCTAAAGGGCTGG 600
    |||||
DB 582 TCATACACCCATGATGATGCTGCGCCCAAGGGGCTCAACACTGTTGGCTAAAGGGCTGG 641
OY 601 AGCTGTACCTCCAGGCCCATACATCCCTTCCAAAGCTCCCAACTACCTACAGAGGAAA 660
    |||||
DB 642 AGCTGTACCTCCAGGCCCATACATCCCTTCCAAAGCTCCCAACTACCTACAGAGGAAA 701
OY 661 CCACGAGTAGAATTCACAGCTTAACACACCCAGACCTGAGCAAAAGTCATGCTGCAGT 720
    |||||
DB 702 CCACGAGTAGAATTCACAGCTTAACACACCCAGACCTGAGCAAAAGTCATGCTGCAGT 761
OY 721 GAAAGGAATGACGGGTGTTCTGTCACCTTTTCTCTAGAGACTGCTAATGAGAGACA 780
    |||||
DB 762 GAAAGGAATGACGGGTGTTCTGTCACCTTTTCTCTAGAGACTGCTAATGAGAGACA 821
OY 781 AACAGGAAATATCGAATTTGACACAGAGCTTGTGAGAGAGGCTGAGATTTCTTCT 840
    |||||
DB 822 AACAGGAAATATCGAATTTGACACAGAGCTTGTGAGAGAGGCTGAGATTTCTTCT 881
OY 841 CCGGAAACAAACACTTATTCAGAGAGAGAGAAATTCGTACTGTGAGAGAGCCCTTCTCT 900
    |||||
DB 882 CCGGAAACAAACACTTATTCAGAGAGAGAGAAATTCGTACTGTGAGAGAGCCCTTCTCT 941
OY 901 ACATCTGGAATGAGACGGTTATGACACACTGATTAATCTGCTCCCTGGCAACCATAT 960
    |||||
DB 942 ACATCTGGAATGAGACGGTTATGACACACTGATTAATCTGCTCCCTGGCAACCATAT 1001
OY 961 TGATCGAATTAATAAGACACTTAATCTATTCATATTCGCTTAGCCCTTGGGGTGGGAG 1020
    |||||
DB 1002 TGATCGAATTAATAAGACACTTAATCTATTCATATTCGCTTAGCCCTTGGGGTGGGAG 1061
OY 1021 AACCTTAGAGTCTCAAGTCAAAAGTGTGCGCTGTGCTGTGAGAGACGCTTCT 1080
    |||||
DB 1062 AACCTTAGAGTCTCAAGTCAAAAGTGTGCGCTGTGCTGTGAGAGACGCTTCT 1121
OY 1081 GCAGGGTGTGAGGCTGACCTTACCTCAGAGTGAGANTGCTCCATGANTACTTTGA 1140
    |||||
DB 1122 GCAGGGTGTGAGGCTGACCTTACCTCAGAGTGAGANTGCTCCATGANTACTTTGA 1181
OY 1141 TGCTGCTTCCCAAGGAATTAATGTATCATCTCTGTGAACAACAGACACTGAACGAGCTT 1200
    |||||
DB 1182 TGCTGCTTCCCAAGGAATTAATGTATCATCTCTGTGAACAACAGACACTGAACGAGCTT 1241
OY 1201 TCTTCTGACCTTGAAGATGCTGTGATTCATCTTGGAGAAATAAGATTAATATCTCT 1260
    |||||
DB 1242 TCTTCTGACCTTGAAGATGCTGTGATTCATCTTGGAGAAATAAGATTAATATCTCT 1301
OY 1261 ATCAGAGACTGACAGAGAGCCCTTCAGTGTGTATTAATTTGAGAAACATCAGAGATAAC 1320
    |||||
DB 1302 ATCAGAGACTGACAGAGAGCCCTTCAGTGTGTATTAATTTGAGAAACATCAGAGATAAC 1361
OY 1321 ATTCT 1324
    |||||
DB 1362 ATTCT 1365

```

RESULT 4
AF283538 1425 bp mRNA linear PRI 16-JAN-2001
LOCUS AF283538
DEFINITION Homo sapiens NIF3L1 protein mRNA, complete cds.

ACCESSION AF283538
 VERSION AF283538.1 GI:12006402
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1425)
 Tascou, S., Uedelhoven, J., Dixkens, C., Neyerla, K., Engel, W. and Burelind, P.
 Isolation and characterization of a novel human gene, NIF3L1, and its mouse ortholog, Nif3l1, highly conserved from bacteria to mammals
 TITLE
 JOURNAL Cyto genet. Cell Genet. 90 (3-4), 330-336 (2000)
 MEDLINE 20573864
 PUBMED 11124544
 REFERENCE 2 (bases 1 to 1425)
 Tascou, S., Burelind, P. and Engel, W.
 Direct Submission
 Submitted (29-JUN-2000) Institute for Human Genetics, University of Goettingen, Heinrich-Dueker Weg 12, Goettingen 37073, Germany
 FEATURES
 Source Location/Qualifiers
 1..1425
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 47..1180
 /codon_start=1
 /product="NIF3L1 protein"
 /protein_id="AAG44846.1"
 /db_xref="GI:12006403"
 Translation="MLSSQVPPVPTTVRPVDSLIGNSSRPMDIKALLSLNDFASIS
 FAEWDNYGLYERPPRPVNTFLTNDEEVMEEVLOKRDILSLPPIPRPKR
 ITWNTKRLVRLAKNRGILSPHAIADAPQVNMALGCACTSPHPSKAP
 IPTENHNHREVNATQDIDKYSVAVKIDGVSFSFSAKRTNEGTSPHLSKAL
 MQVDLFRNKLQYKTELLSLKPLMLTGTGSRICLTDESYSLATMDIRIKHKL
 HIRLAVGRILESOVKYVALCAGSGSSVLCQVEADLYLTGMSHMDLDAASQGINV
 ILCEHSNTERGFLSDRLMDLSLEKNKIILISEEDRDLQVY"
 BASE COUNT 385 a 331 c 314 g 395 t
 ORIGIN
 Query Match 73.3%; Score 1138; DB 9; Length 1425;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 187 AGTCCCAAGACAGCCGGTTGTAGATCCCTGATCTGCATCTGCCCTTCAT 246
 |||||
 70 AGTCCCAAGACAGCCGGTTGTAGATCCCTGATCTGCATCTGCCCTTCAT 129
 |||||
 247 GATTTGAAGGCT 306
 |||||
 130 GATTTGAAGGCT 189
 |||||
 307 TTGGACAATGTTGATTAAGTGTGAACCAAGCCACACATCTGTAATAACACTTT 366
 |||||
 190 TTGGACAATGTTGATTAAGTGTGAACCAAGCCACACATCTGTAATAACACTTT 249
 |||||
 367 CCGTGAACATGACCTGACAGAGAGTGTGAAGAGGTGCTGCAAAAAGAGCAGACT 426
 |||||
 250 CCGTGAACATGACCTGACAGAGAGTGTGAAGAGGTGCTGCAAAAAGAGCAGACT 309
 |||||
 427 CATTTCT 486
 |||||
 310 CATTTCT 369
 |||||
 487 GAAGAGCGCCTGCTGATCCGGCTCTGGAAGACAGAGTGGTATCTACTCTCTCATAC 546
 |||||
 370 GAAGAGCGCCTGCTGATCCGGCTCTGGAAGACAGAGTGGTATCTACTCTCTCATAC 429
 |||||
 547 AGCCATATATGCTGGGCCCCCAGAGGGGTCAACAATGCTGCTAAAGGGCTTGAGCTTG 606
 |||||
 430 AGCCATATATGCTGGGCCCCCAGAGGGGTCAACAATGCTGCTAAAGGGCTTGAGCTTG 489

QY 607 TACCTCCAGGCCCATACATCTCTCCAAAGCTCCCACTACCCCTACAGAGGAACACCGC 666
 |||||
 Db 490 TACCTCCAGGCCCATACATCTCTCCAAAGCTCCCACTACCCCTACAGAGGAACACCGC 549
 |||||
 QY 667 AGTGAATTCACAGCTTAATACACACCAAGACCGGACCAAGATCATCTGCACTGAAGA 726
 |||||
 Db 550 AGTGAATTCACAGCTTAATACACACCAAGACCGGACCAAGATCATCTGCACTGAAGA 609
 |||||
 QY 727 AATTGACGCTGTTCTGTCACCTCTTTCTGCTAGACCTGTAATGAGAACCAACACG 786
 |||||
 Db 610 AATTGACGCTGTTCTGTCACCTCTTTCTGCTAGACCTGTAATGAGAACCAACACG 669
 |||||
 QY 787 GATTATCTGAATTTGACTCAGAGAGCTTGAATGACAGGTGATTTCTTCCCGGA 846
 |||||
 Db 670 GATTATCTGAATTTGACTCAGAGAGCTTGAATGACAGGTGATTTCTTCCCGGA 729
 |||||
 QY 847 CAACACACTTTATCAGAACAGGAAATTTGTCACAGAGAGGCTTTGCTTACATAC 906
 |||||
 Db 730 CAACACACTTTATCAGAACAGGAAATTTGTCACAGAGAGGCTTTGCTTACATAC 789
 |||||
 QY 907 TGGATGGAAGCGTTATGACACACTGATGAATCTGTCCTCGCAACCATGATGATCG 966
 |||||
 Db 790 TGGATGGAAGCGTTATGACACACTGATGAATCTGTCCTCGCAACCATGATGATCG 849
 |||||
 QY 967 AATTAAGACACCTTAACCTATCTCATATTCGCTTAAGCCCTTGGGGGGAACCTT 1026
 |||||
 Db 850 AATTAAGACACCTTAACCTATCTCATATTCGCTTAAGCCCTTGGGGGGAACCTT 909
 |||||
 QY 1027 AGAGTCTCAAGTCAAAAGTGTGAGCCTGTGTGCTGTGTGAGGAGCAGCTTTCGAGG 1086
 |||||
 Db 910 AGAGTCTCAAGTCAAAAGTGTGAGCCTGTGTGCTGTGTGAGGAGCAGCTTTCGAGG 969
 |||||
 QY 1087 TGTGAGGCTGACCTTTACCTCAACAGTGAATGTCCTCATGATCTTTGATCTGC 1146
 |||||
 Db 970 TGTGAGGCTGACCTTTACCTCAACAGTGAATGTCCTCATGATCTTTGATCTGC 1029
 |||||
 QY 1147 TTCCCAAGAAATTAATGTCATCTCTGTGAACACACACACAGAGAGGCTTTCTTC 1206
 |||||
 Db 1030 TTCCCAAGAAATTAATGTCATCTCTGTGAACACACACACAGAGAGGCTTTCTTC 1089
 |||||
 QY 1207 TGACCTTGAGATATGCTGATTCATCTGATGGAATTAAGTAATTAATCTATCAGA 1266
 |||||
 Db 1090 TGACCTTGAGATATGCTGATTCATCTGATGGAATTAAGTAATTAATCTATCAGA 1149
 |||||
 QY 1267 GACTGACAGGAGCCCTCTTCAAGTGTGTAATTAATGCAAGAAACATCAGATTAACATTC 1324
 |||||
 Db 1150 GACTGACAGGAGCCCTCTTCAAGTGTGTAATTAATGCAAGAAACATCAGATTAACATTC 1207
 |||||
 RESULT 5
 LOCUS BC007654 1440 bp mRNA linear PRI 12-JUL-2001
 DEFINITION Homo sapiens, similar to NIF3 (Ngi) interacting factor 3, S. Pombe
 ACCESSION BC007654
 VERSION BC007654.1 GI:14043316
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1440)
 Strausberg, R.
 Direct Submission
 Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC/DCPD/FTP

CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgcnhgrl.nih.gov
 Stevehenko, Y., Wetherby, R.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, O.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stantilop, S., Thomas, P.J.,
 Tjongson, E.E., Touchman, J.W., Tsurgoun, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 5 Row: 1 Column: 12
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10197631.
 Location/Qualifiers

FEATURES

source

COS

```

1..1440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:2595 IMAGE:3346099"
/tissue_type="Skin, melanotic melanoma."
/clone_id="N1H-MG-20"
/lab_host="DH10B-R"
/notes="vector: pOTB7"
37..1170
/codon_start=1
/product="Similar to NIF3 (Nng1 interacting factor 3,
s.pombe homolog)-like 1"
/protein_id="AA07654.1"
/db_xref="GI:14043317"
/translocation="MUSCVRPVTYRYEYDLSLNCSSRSFMDIKALLSLUNDFASIS
FASMDVGLVLEPSPHPTVNTLFLINDLIEVWEVLQKADILSYHPERPMKR
ITWNTKERVIALENRNGIYSPHTAYDAPQSVNMMLAKGLGACTSRIPHSKAR
YPTGNHREYFNNTODLDKVMASVAGIGDSVTSFSARNGEORINLNCOKL
MOVVDPLSRKOLYOKREILSLKRPILLHGMGLCTLDSSVLSIATWIDIRLRLXIS
HIRALAGRTLSQVKYVALCAGSSSVYQGVADYLYTGEMSHHTDLAASGIVN
ILCESHTERGFJSDLRMDLSDHKNRINILSETDPLQV"
BASE COUNT      407 a      329 c      313 g      391 t
ORIGIN

```

Query Match 73.3% Score 1138; DB 9; Length 1440;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 187 AGTCCCGACAGACAGTCCGGTTGTAGATCCCTGATCTGCATATCTCCGTTCCCTCAT 246
    |||||||
Db 60 AGTCCCGACAGACAGTCCGGTTGTAGATCCCTGATCTGCATATCTCCGTTCCCTCAT 119
    |||||||

QY 247 GGAATTGAAGGCTCTCTTCTTCTTGAATGACTTTGATCCCTCTGCTTGTCTGAGAG 306
    |||||||
Db 120 GGAATTGAAGGCTCTCTTCTTCTTGAATGACTTTGATCCCTCTGCTTGTCTGAGAG 179
    |||||||

QY 307 TTGGGACAAATGTTGATTAAGTGAACCAAGCCACACATCTGTAATAATACACTCTT 366
    |||||||
Db 180 TTGGGACAAATGTTGATTAAGTGAACCAAGCCACACATCTGTAATAATACACTCTT 239
    |||||||

QY 367 CTTGACCAATGACTGACTGAGAAAGTATGAGAGAGTGTCTGCAAAAAGAGGAGACT 426
    |||||||
Db 240 CTTGACCAATGACTGACTGAGAAAGTATGAGAGAGTGTCTGCAAAAAGAGGAGACT 299
    |||||||

QY 427 CATTTCTCTACCATCCGCTATCTTCCGACCATGAAAGGCAATTAACCTGGAACACATG 486
    |||||||
Db 300 CATTTCTCTACCATCCGCTATCTTCCGACCATGAAAGGCAATTAACCTGGAACACATG 359
    |||||||

QY 487 GAAGAGAGGCTGTGATCCGGGCTTGAGAAACAGAGTGGTATCTACTCTCTCCTACAC 546
    |||||||
Db 360 GAAGAGAGGCTGTGATCCGGGCTTGAGAAACAGAGTGGTATCTACTCTCTCCTACAC 419
    |||||||

```

```

QY 547 AGCCTATGATGCTGCGCCCGCCAGGGCGTCAACACATGTTGGCTAAAGGCTTGGAGCTTG 606
    |||||||
Db 420 AGCCTATGATGCTGCGCCCGCCAGGGCGTCAACACATGTTGGCTAAAGGCTTGGAGCTTG 479
    |||||||

QY 607 TACTCCAGGCCCCATACATCTCTTCCAAAGTCCCACTACCTCTACAGAGAAACACCG 666
    |||||||
Db 480 TACTCCAGGCCCCATACATCTCTTCCAAAGTCCCACTACCTCTACAGAGAAACACCG 539
    |||||||

QY 667 AGTAGATTCAGCTTAAGTACACCAAGACCTGGACAAAGTCAATGCTGAGTGAAGAAG 726
    |||||||
Db 540 AGTAGATTCAGCTTAAGTACACCAAGACCTGGACAAAGTCAATGCTGAGTGAAGAAG 599
    |||||||

QY 727 AATTGAGGTTTCTGTCACTTCTTTTCTGTCTAGGACTGTGAATGAGAAACAACAGC 786
    |||||||
Db 600 AATTGAGGTTTCTGTCACTTCTTTTCTGTCTAGGACTGTGAATGAGAAACAACAGC 659
    |||||||

QY 787 GATTAAATCTGATTTGTACTCAGAAAGGCTTTGATGACAGTGTAGATTTCTTCCGGAA 846
    |||||||
Db 660 GATTAAATCTGATTTGTACTCAGAAAGGCTTTGATGACAGTGTAGATTTCTTCCGGAA 719
    |||||||

QY 847 CAACCAACTTATCAGAAAGGAAATTCGTCACTGAGAAAGCTTGTCTTACATAC 906
    |||||||
Db 720 CAACCAACTTATCAGAAAGGAAATTCGTCACTGAGAAAGCTTGTCTTACATAC 779
    |||||||

QY 907 TGAATGAGGAGGTTATGACACACTGATGAATCTGTCTCCCTGGCAACATGATGATCG 966
    |||||||
Db 780 TGAATGAGGAGGTTATGACACACTGATGAATCTGTCTCCCTGGCAACATGATGATCG 839
    |||||||

QY 967 AATAAAGACACCTAAACATCTCATATTCGTTAGCCTTTGGGCTGGGGAACACTT 1026
    |||||||
Db 840 AATAAAGACACCTAAACATCTCATATTCGTTAGCCTTTGGGCTGGGGAACACTT 899
    |||||||

QY 1027 AGACTCTCAAGTCAAGTCGTCGCGCCGTGTGCTGTCTGTGGAGAGACGCTTCTGACAGG 1086
    |||||||
Db 900 AGACTCTCAAGTCAAGTCGTCGCGCCGTGTGCTGTCTGTGGAGAGACGCTTCTGACAGG 959
    |||||||

QY 1087 TGTGAGGCTGACCTTAACTCAGAGGTGAGATGTCCATCATGATCTTTGGATGCTGC 1146
    |||||||
Db 960 TGTGAGGCTGACCTTAACTCAGAGGTGAGATGTCCATCATGATCTTTGGATGCTGC 1019
    |||||||

QY 1147 TTCCCAAGGAATTAATGTATCTCTCTGTGAACAAGCAACCTGAAGAGCTTCTTTC 1206
    |||||||
Db 1020 TTCCCAAGGAATTAATGTATCTCTCTGTGAACAAGCAACCTGAAGAGCTTCTTTC 1079
    |||||||

QY 1207 TGACCTCGAGATATGCTGATCTCTCACTTGGAGATTAAGTAAATATATCTATGAGA 1266
    |||||||
Db 1080 TGACCTCGAGATATGCTGATCTCTCACTTGGAGATTAAGTAAATATATCTATGAGA 1139
    |||||||

QY 1267 GACTGACAGGACCCCTTTCAGGTGTATTAATTCAGAAACATCAGATTAACACATTC 1324
    |||||||
Db 1140 GACTGACAGGACCCCTTTCAGGTGTATTAATTCAGAAACATCAGATTAACACATTC 1197
    |||||||

```

RESULT 6
 AF060513 1387 bp mRNA linear PRI 02-JAN-2001
 LOCUS Homo sapiens clone 016d06 My018 protein mRNA, complete cds.
 DEFINITION AF060513
 ACCESSION AF060513
 VERSION AF060513.1 GI:12001975
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 1387)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Mao, Y.M., Xie, Y., Huang, X.Y., Ying, K. and Dai, J.L.
 TITLE Direct Submission
 JOURNAL Submitted (20-APR-1998) Institute of Genetics, School of Life
 Science, Fudan University, 220 Handan Rd., Shanghai 200433,
 P. R. China

FEATURES
 source Location/Qualifiers
 1..1387

```

/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="016806"
/tissue_type="brain"
/dev_stage="fetus"
82..1134
/codon_start=1
/product="My018 protein"
/protein_id="MAG43131.1"
/db_xref="GI:12001976"
/translacion="MDLKALSLNDFASLSFAESMDNVLVEPSPHTVNTLEFLN
DLFEVMEVILOKKADLILSYHPPIRPMKRITMWTKEKRLVIALENRGISPTA
YDAAPGVNMMILAKGIGACTSRPIHPSKAPNYPEGNHREPVNVTODIDKMSAVK
GIDGVSYSFSAFTGNEBOTRINLCTOKALMOYVDELIRNKOLOKXETELSLERPL
LHTGMKRLCTLDSESVSLATMIDIRIKRLKLSHRLALGRTLESQVKKVALDAGSGS
SVLQGEADLYLTGEMSHDPTLDAASOGINVILCEHSNERGFLSDLRLMDLSHLENK
INITLSETRDPLQVV"
BASE COUNT      378 a      323 c      309 g      377 t
ORIGIN

Query Match      73.2% Score 1137; DB 9; Length 1387;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 GTCCCCACGACAGTCCGTTTGATGATTCCTGATGCAATTCCTCCGTCCTTCATG 247
    |||||
DB 25 GTCCCCACGACAGTCCGTTTGATGATTCCTGATGCAATTCCTCCGTCCTTCATG 84
    |||||

QY 248 GATTTGAAGGCTCTCTTCTTCTTCTTGAATGACTTTGACTCCCTCTGCTGAGAGT 307
    |||||
DB 85 GATTTGAAGGCTCTCTTCTTCTTCTTGAATGACTTTGACTCCCTCTGCTGAGAGT 144
    |||||

QY 308 TGGGACAAATGTTGATTAAGTGGAACCAAGCCACATCTGTAAGAACACTCTTC 367
    |||||
DB 145 TGGGACAAATGTTGATTAAGTGGAACCAAGCCACATCTGTAAGAACACTCTTC 204
    |||||

QY 368 CTGACCAATGACTGACTGAGAAAGTATGAGAGAGTGTCTGCAAAAAGAGGACACTC 427
    |||||
DB 205 CTGACCAATGACTGACTGAGAAAGTATGAGAGAGTGTCTGCAAAAAGAGGACACTC 264
    |||||

QY 428 ATTCTCTCTACCATCCGCTATCTTCCGACCATGAAGGCGCTTAACCTGGAACACATGG 487
    |||||
DB 265 ATTCTCTCTACCATCCGCTATCTTCCGACCATGAAGGCGCTTAACCTGGAACACATGG 324
    |||||

QY 488 AAGGACGCGCTGGTGTATCCGGGCTCTGGAAGACAGAGTGGTATCTCTCCCTCATACA 547
    |||||
DB 325 AAGGACGCGCTGGTGTATCCGGGCTCTGGAAGACAGAGTGGTATCTCTCCCTCATACA 384
    |||||

QY 548 GCCCTATGATGCTGCGGCCCGACAGGCGCTCAACACTGTTGGCTAAAGGGCTTGGAGCTTGT 607
    |||||
DB 385 GCCCTATGATGCTGCGGCCCGACAGGCGCTCAACACTGTTGGCTAAAGGGCTTGGAGCTTGT 444
    |||||

QY 608 ACCTCCAGGCGCTATACATCTTCCAAAGCTCCCACTACCTCTACAGAGGAAACACACCGA 667
    |||||
DB 445 ACCTCCAGGCGCTATACATCTTCCAAAGCTCCCACTACCTCTACAGAGGAAACACACCGA 504
    |||||

QY 668 GTGAAATTCAGCTTAACACCAAGACCTGGACAAAGTCTATGTCTGAGTGAAGAAGA 727
    |||||
DB 505 GTGAAATTCAGCTTAACACCAAGACCTGGACAAAGTCTATGTCTGAGTGAAGAAGA 564
    |||||

QY 728 ATTGAGGCTGTTCTGTCACTTCTTCTTCTGATGAGACTGATGATGAGAAACACAGG 787
    |||||
DB 565 ATTGAGGCTGTTCTGTCACTTCTTCTTCTGATGAGACTGATGATGAGAAACACAGG 624
    |||||

QY 788 ATTAATCTGAATGTACTCAGAAAGCTTTGATGACAGGTGATGATTTCTTCCCGGAC 847
    |||||
DB 625 ATTAATCTGAATGTACTCAGAAAGCTTTGATGACAGGTGATGATTTCTTCCCGGAC 684
    |||||

QY 848 AAACAATTTATGAGAAAGCGAAATTCCTCACTGAGAAAGCTTTGCTTCACTACT 907
    |||||
DB 685 AAACAATTTATGAGAAAGCGAAATTCCTCACTGAGAAAGCTTTGCTTCACTACT 744
    |||||

```

```

QY 908 GGAATGGAGAGGTTATGACACATGATGATTCCTGCTCCCTGGCAACCATGATGATGCA 967
    |||||
DB 745 GGAATGGAGAGGTTATGACACATGATGATTCCTGCTCCCTGGCAACCATGATGATGCA 804
    |||||

QY 968 ATAAAAAGACACCTAAAACTATCTCATATTCGTTAGCCCTTGGGGTGGAGAACCTTA 1027
    |||||
DB 805 ATAAAAAGACACCTAAAACTATCTCATATTCGTTAGCCCTTGGGGTGGAGAACCTTA 864
    |||||

QY 1028 GACTCTCAAGTCAAAAGCTGCGGCCCTGTGCTGCTTCTGGAGACAGCTTCTGCAAGGT 1087
    |||||
DB 865 GACTCTCAAGTCAAAAGCTGCGGCCCTGTGCTGCTTCTGGAGACAGCTTCTGCAAGGT 924
    |||||

QY 1088 GTTGAGGCTACCTTTACCTACACAGGAGATGATGCCATGATGATTTGATGATGCT 1147
    |||||
DB 925 GTTGAGGCTACCTTTACCTACACAGGAGATGATGCCATGATGATTTGATGATGCT 984
    |||||

QY 1148 TCCCAAGGAATTAATGTATCTCTGTGAACACAGCAACACTGAAAGAGCTTCTTCT 1207
    |||||
DB 985 TCCCAAGGAATTAATGTATCTCTGTGAACACAGCAACACTGAAAGAGCTTCTTCT 1044
    |||||

QY 1208 GACCTTCGATATATGCTGATTCCTGATGATGATGATGATGATGATGATGATGATGAT 1267
    |||||
DB 1045 GACCTTCGATATATGCTGATTCCTGATGATGATGATGATGATGATGATGATGATGAT 1104
    |||||

QY 1268 ACTGACAGGAGACCTTCTCAGTGTATTAATTCAGAAATCAGATGATGATGATGATGAT 1324
    |||||
DB 1105 ACTGACAGGAGACCTTCTCAGTGTATTAATTCAGAAATCAGATGATGATGATGATGAT 1161
    |||||

RESULT 7
AX119075
LOCUS
DEFINITION
Sequence 239 from Patent WO0129221.
AX119075
ACCESSION
AX119075.1 GI:14036029
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ORGANISM

REFERENCE
1 Konkili, D.C. and Yee, D.P.
Proteins and polynucleotides encoding them
Patent: WO 0129221-A 239 26-APR-2001;
Zymogenetics, Inc. (US)
JOURNAL
LOCATION/Qualifiers
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..1053
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAC38611.1"
/db_xref="GI:14036030"
/translacion="MDLKALSLNDFASLSFAESMDNVLVEPSPHTVNTLEFLN
DLFEVMEVILOKKADLILSYHPPIRPMKRITMWTKEKRLVIALENRGISPTA
YDAAPGVNMMILAKGIGACTSRPIHPSKAPNYPEGNHREPVNVTODIDKMSAVK
GIDGVSYSFSAFTGNEBOTRINLCTOKALMOYVDELIRNKOLOKXETELSLERPL
LHTGMKRLCTLDSESVSLATMIDIRIKRLKLSHRLALGRTLESQVKKVALDAGSGS
SVLQGEADLYLTGEMSHDPTLDAASOGINVILCEHSNERGFLSDLRLMDLSHLENK
INITLSETRDPLQVV"
BASE COUNT      279 a      253 c      247 g      274 t
ORIGIN

Query Match      67.8% Score 1053; DB 6; Length 1053;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 ATGATTTGAAGGCTCTCTTCTTCTTCTTGAAGACTTTCATCCCTCTGTTGCTGAG 304
    |||||
DB 1 ATGATTTGAAGGCTCTCTTCTTCTTCTTGAAGACTTTCATCCCTCTGTTGCTGAG 60
    |||||

```


QY 305 AGTTGGACAATGTGGATTACTGCTGGGAACCAAGCCACACACTGTAATAATACACTC 364
 |||||
 Db 61 AGTTGGACAATGTGGATTACTGCTGGGAACCAAGCCACACACTGTAATAATACACTC 120
 |||||
 QY 365 TTCCCTGACCAATACCTGATGAGAAAGTATGAGAGAGTGTGCAAAAGAGGCGAC 424
 |||||
 Db 121 TTCCCTGACCAATACCTGATGAGAAAGTATGAGAGAGTGTGCAAAAGAGGCGAC 180
 |||||
 QY 425 CTGATTCCTCTACCAATCCGCTATCTCCGACCCCTGAAAGGCGATTAACCTGGAACACA 484
 |||||
 Db 181 CTGATTCCTCTACCAATCCGCTATCTCCGACCCCTGAAAGGCGATTAACCTGGAACACA 240
 |||||
 QY 485 TGAAGAGAGCGCTGTGATCCGGGCTCTGAGAACAGAGTGGATATCTCTCTCAT 544
 |||||
 Db 241 TGAAGAGAGCGCTGTGATCCGGGCTCTGAGAACAGAGTGGATATCTCTCTCAT 300
 |||||
 QY 545 ACAGCCATATGATGCTGGGCGCCGAGGCGTCAACAACTGGTGTAAAGGCGTTGGAGCT 604
 |||||
 Db 301 ACAGCCATATGATGCTGGGCGCCGAGGCGTCAACAACTGGTGTAAAGGCGTTGGAGCT 360
 |||||
 QY 605 TGTACCTCCAGGCCATACATCTCTCCAAAGCTCCCAACTACCTACAGAGGGAACAC 664
 |||||
 Db 361 TGTACCTCCAGGCCATACATCTCTCCAAAGCTCCCAACTACCTACAGAGGGAACAC 420
 |||||
 QY 665 CGAGTAGAATTCACAGTTAACTACACCCAGAGCTGGACAAAGTCTGCACTGATA 724
 |||||
 Db 421 CGAGTAGAATTCACAGTTAACTACACCCAGAGCTGGACAAAGTCTGCACTGATA 480
 |||||
 QY 725 GGATTCAGGCTGTTCTGTACACTCTTTCTGCTGAGAGCTGGTAAAGAGAACAAACA 784
 |||||
 Db 481 GGATTCAGGCTGTTCTGTACACTCTTTCTGCTGAGAGCTGGTAAAGAGAACAAACA 540
 |||||
 QY 785 CGGATTAATCTGAATTTACTCAGAGAGCTTTGATGAGAGTGTAGATTTCTTCCCGG 844
 |||||
 Db 541 CGGATTAATCTGAATTTACTCAGAGAGCTTTGATGAGAGTGTAGATTTCTTCCCGG 600
 |||||
 QY 845 AACAAACAACTTTATCAGAAAGACGGAATTTCTGATCAGAGAGCCCTTCTCTACAT 904
 |||||
 Db 601 AACAAACAACTTTATCAGAAAGACGGAATTTCTGATCAGAGAGCCCTTCTCTACAT 660
 |||||
 QY 905 ACTGGAATGGAGCGGTTATGACACACTGATGAATCTGCTCCCTGGCAACCATGATGAT 964
 |||||
 Db 661 ACTGGAATGGAGCGGTTATGACACACTGATGAATCTGCTCCCTGGCAACCATGATGAT 720
 |||||
 QY 965 CGAATTAACAAACACTTAAACATATCTATATTCGCTTAGCCCTTGGGGTGGGAGAAC 1024
 |||||
 Db 721 CGAATTAACAAACACTTAAACATATCTATATTCGCTTAGCCCTTGGGGTGGGAGAAC 780
 |||||
 QY 1025 TTAGAGTCTCAAGTCAAGAAAGTGGGCGCTGTGCTGTGCTGGGAGAGAGGTTCTGAG 1084
 |||||
 Db 781 TTAGAGTCTCAAGTCAAGAAAGTGGGCGCTGTGCTGTGCTGGGAGAGAGGTTCTGAG 840
 |||||
 QY 1085 GGTGTGAGGCTGACCTTTACTCCTCAGAGTGAATGTCATCATGATTAATTTGGATGCT 1144
 |||||
 Db 841 GGTGTGAGGCTGACCTTTACTCCTCAGAGTGAATGTCATCATGATTAATTTGGATGCT 900
 |||||
 QY 1145 GCTTCCCAAGGAATTAATGTCATCTCTGTGAAACACAGCAACTGAAAGAGGCTTTCTT 1204
 |||||
 Db 901 GCTTCCCAAGGAATTAATGTCATCTCTGTGAAACACAGCAACTGAAAGAGGCTTTCTT 960
 |||||
 QY 1205 TCTGACCTTCAGATATGCTGATTTCTCACTTGGAGATTAAGATTAATTAATCTATACA 1264
 |||||
 Db 961 TCTGACCTTCAGATATGCTGATTTCTCACTTGGAGATTAAGATTAATTAATCTATACA 1020
 |||||
 QY 1265 GAGACTGACAGGAGCCCTCTCAGGTGGTATTA 1297
 |||||
 Db 1021 GAGACTGACAGGAGCCCTCTCAGGTGGTATTA 1053
 |||||

RESULT 8
 HSM805522
 LOCUS
 DEFINITION Homo sapiens mRNA; cDNA DKFZp762L015 (from clone DKFZp762L015).

ACCESSION AL834430
 VERSION AL834430.1 GI:21740154
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1353)
 REFERENCE
 AUTHORS Bloembergen, H., Boecker, M., Brandt, P., Mewes, H.W., Well, B. and
 Wiemann, S.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUL-2002) 1. D-85764 Neuherberg, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 Brunschwelg/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFZp762L015) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcententrum, Neuherberg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://mips.gsf.de/proj/cDNA/>.
 Location/Qualifiers
 FEATURES
 source
 1. 1353
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp762L015"
 /tissue_type="melanoma (Melo cell line)"
 /clone_lib="762 (synonym: hmel2). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 polyA_signal
 1228. 1233
 BASE COUNT 457 a 283 c 280 g 333 t
 ORIGIN
 polyA_site
 1228. 1233
 Query Match 66.1%; Score 1027; DB 9; Length 1353;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 298 TGCTGAGAGTTGGCAATGTTGATTAATCTGTTGGAACCAAGCCACATCTGTA 357
 |||||
 Db 1 TGCTGAGAGTTGGCAATGTTGATTAATCTGTTGGAACCAAGCCACATCTGTA 60
 |||||
 QY 358 TACACTCTTCTGACCAATGACACCTGACAGAGAGTGTGAGAGAGGTCGCAAAAGA 417
 |||||
 Db 61 TACACTCTTCTGACCAATGACACCTGACAGAGAGTGTGAGAGAGGTCGCAAAAGA 120
 |||||
 QY 418 GGCAGACCTCATTTCTCTCTACATCCGCTATCTTCCGACCATGAAGCGCATTAACCTG 477
 |||||
 Db 121 GGCAGACCTCATTTCTCTCTACATCCGCTATCTTCCGACCATGAAGCGCATTAACCTG 180
 |||||
 QY 478 GAACACATGAGAGAGCGCTGATGATCCGGGCTCTGGAAGACAGAGTCGATATCTAC 537
 |||||
 Db 181 GAACACATGAGAGAGCGCTGATGATCCGGGCTCTGGAAGACAGAGTCGATATCTAC 240
 |||||
 QY 538 TCCTATACAGCCTATGATGCTGCGCCCGGAGGCGCTCAACAACGTTGGCTAAAGGCT 597
 |||||
 Db 241 TCCTATACAGCCTATGATGCTGCGCCCGGAGGCGCTCAACAACGTTGGCTAAAGGCT 300
 |||||
 QY 598 TGGAGCTGTACCTCCAGGCCATACATCTCTCCAAAGCTCCCACTACCTACAGAGG 657
 |||||
 Db 301 TGGAGCTGTACCTCCAGGCCATACATCTCTCCAAAGCTCCCACTACCTACAGAGG 360
 |||||
 QY 658 AAACACGAGTAGAATTCACAGTTAACTACACCCAGAGCTGACCAAAAGTATCTGTC 717
 |||||
 Db 361 AAACACGAGTAGAATTCACAGTTAACTACACCCAGAGCTGACCAAAAGTATCTGTC 420
 |||||
 QY 718 AGTGAAGGAATGACGCTGTTCTGCTCACTCTTTCTTCTGCTAGAGCTGTAATGAGA 777
 |||||
 Db 421 AGTGAAGGAATGACGCTGTTCTGCTCACTCTTTCTTCTGCTAGAGCTGTAATGAGA 480
 |||||

QY	778	ACAAACGCGTTATCTGATTCGTACTCAGAAAGCGCTTGGATGCAAGTGGTGAATTTTCT	837
Db	481	ACAAACGCGTTATCTGATTCGTACTCAGAAAGCGCTTGGATGCAAGTGGTGAATTTTCT	540
QY	838	TTCCCGGAACAAACCTTTATCTGAAAGACGGAATTCGTACCTGGAGAACCTTTGCT	897
Db	541	TTCCCGGAACAAACCTTTATCTGAAAGACGGAATTCGTACCTGGAGAACCTTTGCT	600
QY	898	TCATCATCTGGAATGGACGCGTTATGCACACTGATGAATCTGTCTCCCTGGCAACCAT	957
Db	601	TCATCATCTGGAATGGACGCGTTATGCACACTGATGAATCTGTCTCCCTGGCAACCAT	660
QY	958	GATTGATCGAATAAAAAGACACCTAAACCTATCTCATATTCGCTTAGCGCTTGGGGTGGG	1017
Db	661	GATTGATCGAATAAAAAGACACCTAAACCTATCTCATATTCGCTTAGCGCTTGGGGTGGG	720
QY	1018	GAGAACCTTAGAGTCTCAAGTCAAAAGTCTGGCCCTGTGCTGTTCTGGGACACAGT	1077
Db	721	GAGAACCTTAGAGTCTCAAGTCAAAAGTCTGGCCCTGTGCTGTTCTGGGACACAGT	780
QY	1078	TCCTCAGGGGTCTTAGAGCTGTACCTTTACCTGCACAGGTGAGATGTCCATCATGATCTTT	1137
Db	781	TCCTCAGGGGTCTTAGAGCTGTACCTTTACCTGCACAGGTGAGATGTCCATCATGATCTTT	840
QY	1138	GGATGCTGCTTCCCAAGGAATTAATGTGATCCTGTGACACACAGCAACATGAAACGAG	1197
Db	841	GGATGCTGCTTCCCAAGGAATTAATGTGATCCTGTGACACACAGCAACATGAAACGAG	900
QY	1198	CTTCTCTTCTGACCTCGAGATATGCTGGATTCCTGCACCTGGAGAAATGAATATATAT	1257
Db	901	CTTCTCTTCTGACCTCGAGATATGCTGGATTCCTGCACCTGGAGAAATGAATATATAT	960
QY	1258	CCATATCAGAGACTGACAGGACCCCTCTCAGGTGATTAATTCAGAAACATCAGAGTAA	1317
Db	961	CCATATCAGAGACTGACAGGACCCCTCTCAGGTGATTAATTCAGAAACATCAGAGTAA	1020
QY	1318	CACATTC 1324	
Db	1021	CACATTC 1027	
RESULT 9			
LOCUS	AF182416	1579 bp	linear mRNA PRI 20-SEP-2000
DEFINITION	Homo sapiens MD5015 (MD5015) mRNA, complete cds.		
ACCESSION	AF182416		
VERSION	AF182416.1	GI:10197631	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 1579)		
TITLE	Huang, C., Qian, B., Tu, Y., Gu, W., Wang, Y., Han, Z. and Chen, Z.		
JOURNAL	Novel genes expressed in hematopoietic stem/progenitor cells from Myelodysplastic Syndromes patient		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 1579)		
TITLE	Huang, C., Qian, B., Tu, Y., Gu, W., Wang, Y., Han, Z. and Chen, Z.		
JOURNAL	Direct Submission		
FEATURES	Submitted (02-SEP-1999) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, People's Republic of China		
source	Location/Qualifiers		
	1..1579		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/cell_type="hematopoietic stem/progenitor cells"		
	/note="Isolated from a Myelodysplastic Syndrome patient"		
	1..1579		
	/gene="MD5015"		
	244..1299		
CDS			

[illegible]

Db 780 ACAAACAGGATTAATCTGATTTGATCTACAGAGGCTTTGATGAGGTGATGATTTTCT 839

QY 838 TTCCCGAACAACAACCTTTATTCAGAGAGGAAATTCCTGACTGAGAGACCTTCTCT 897

Db 840 TTCCCGAACAACAACCTTTATTCAGAGAGGAAATTCCTGACTGAGAGACCTTCTCT 899

QY 898 TCTACATAGTGAATGGAGCGTTATGACACTGGATGAATCTGTCCTCCGCAACCAT 957

Db 900 TCTACATAGTGAATGGAGCGTTATGACACTGGATGAATCTGTCCTCCGCAACCAT 959

QY 958 GATTGATCGAATAAAGACACCTTAAACATATCTCATTTGCGTTAGCCCTTGGGGTGG 1017

Db 960 GATTGATCGAATAAAGACACCTTAAACATATCTCATTTGCGTTAGCCCTTGGGGTGG 1019

QY 1018 GAGAAGCTTGAAGTCTCAAGTCAAGTCTGGCCCTGCTGCTGTTCTGGAGACAGCT 1077

Db 1020 GAGAAGCTTGAAGTCTCAAGTCAAGTCTGGCCCTGCTGCTGTTCTGGAGACAGCT 1079

QY 1078 TCTGCAAGGGTGTGAGGCTGACCTTTACCTGACAGGTGAGATGTCCATCATGATCTTT 1137

Db 1080 TCTGCAAGGGTGTGAGGCTGACCTTTACCTGACAGGTGAGATGTCCATCATGATCTTT 1139

QY 1138 GGATGCTGCTTCCCAAGGAATAAATGTCATCTCTGTGACACAGCACTGAAAGCAGG 1197

Db 1140 GGATGCTGCTTCCCAAGGAATAAATGTCATCTCTGTGACACAGCACTGAAAGCAGG 1199

QY 1198 CTTTCTTCTGACCTTCGAGATATGCTGATCTCTGATCTGAGATAAGATAATATTTAT 1257

Db 1200 CTTTCTTCTGACCTTCGAGATATGCTGATCTCTGATCTGAGATAAGATAATATTTAT 1259

QY 1258 CCTATCAGAGACTGACAGGAGCCCTCTGAGGTGGTATTAATTCGAGAAACATCAGATTA 1317

Db 1260 CCTATCAGAGACTGACAGGAGCCCTCTGAGGTGGTATTAATTCGAGAAACATCAGATTA 1319

QY 1318 CACATTC 1324

Db 1320 CACATTC 1326

RESULT 10

BD149184 796 bp DNA linear PAT 17-JAN-2003

LOCUS BD149184

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD149184.1 GI:27854942

VERSION BD149184.1

KEYWORDS JP 2002191363-A/4027.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 796)

AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002191363-A 4027 09-JUL-2002;

COMMENT HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/4027

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU PI SAITO,

PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI NAGAI,TETSUJI OTSUKI

PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10,C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH key

FT source 1..796

FT Location/Qualifiers

FEATURES

source 1..796

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 192 a 204 c 202 g 195 t 3 others

ORIGIN

Query Match 36.5%; Score 567; DB 6; Length 796;

Best Local Similarity 100.0%; Pred. No. 2.8e-269;

Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGATTTGTTATCTTGCTGCTGACAGAGACAGACAGAGAGAGATTGGTCAAGAACTGC 60

Db 27 GTGATTTGTTATCTTGCTGCTGACAGAGACAGACAGAGAGAGATTGGTCAAGAACTGC 86

QY 61 CCTGCGGACACAGACAGACAGGCTAGTGGAGACAGGGGCTCTGACTCAGACTTAACCTG 120

Db 87 CCTGCGGACACAGACAGACAGGCTAGTGGAGACAGGGGCTCTGACTCAGACTTAACCTG 146

QY 121 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

Db 147 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 206

QY 181 TAGATGATCCCCACAGACAGTCCGTTGTGATGATCCCTGATCCGATTTCTCCCTTC 240

Db 207 TAGATGATCCCCACAGACAGTCCGTTGTGATGATCCCTGATCCGATTTCTCCCTTC 266

QY 241 CTTTCATGATTTGAAGGCT 300

Db 267 CTTTCATGATTTGAAGGCT 326

QY 301 TGAGAGTTGGGACAAATGTTGATTAAGTGTGAGAACCAAGCCACACATCTGTAATAC 360

Db 327 TGAGAGTTGGGACAAATGTTGATTAAGTGTGAGAACCAAGCCACACATCTGTAATAC 386

QY 361 ACTCTCCGACCAATGACCTGACAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAG 420

Db 387 ACTCTCTGACCAATGACCTGACAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAG 446

QY 421 AGACCTCATTTCTCTCTCAATCCGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480

Db 447 AGACCTCATTTCTCTCTCAATCCGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 506

QY 481 CACATGAGAGAGGCGCTGATGATCCGGGCTCTGAGAGACAGAGTCTGATCTACTCTCC 540

Db 507 CACATGAGAGAGGCGCTGATGATCCGGGCTCTGAGAGACAGAGTCTGATCTACTCTCC 566

QY 541 TCATACAGCTTATGATGCTGCGGCCCA 567

Db 567 TCATACAGCTTATGATGCTGCGGCCCA 593

RESULT 11

AC037455 170586 bp DNA linear HTG 30-AUG-2001

LOCUS AC037455/c

DEFINITION Homo sapiens chromosome 02 clone RP11-422L5, WORKING DRAFT

ACCESSION AC037455

VERSION AC037455.5 GI:9887641

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 170586)

AUTHORS Smith,D.R.

TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome

JOURNAL Sequence Data

REFERENCE Unpublished

AUTHORS Smith,D.R.

TITLE Direct Submission

JOURNAL Submitted (09-APR-2000) Genome Therapeutics Corporation, 100 Beaver

COMMENT

Street, Waltham, MA 02453, USA
On Aug 24, 2000 this sequence version replaced gi:8569072.

Genome Center

Center: Genome Therapeutics Corporation

Center code: GTC

Web site: <http://www.genomecorp.com/>

Contact: gtc-seqcenter@genomecorp.com

Project Information

Center project name: hg215

Summary Statistics

Sequencing vector: N/A

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 990315

Consensus quality: 147001 bases at least Q40

Consensus quality: 157396 bases at least Q30

Consensus quality: 159713 bases at least Q20

Insert size: 168303; sum-of-contigs

Quality coverage: 3.8x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1171: contig of 1171 bp in length
1172 1271: gap of unknown length
1272 2319: contig of 1048 bp in length
2320 2419: gap of unknown length
2420 3793: contig of 1374 bp in length
3794 3893: gap of unknown length
3894 5414: contig of 1521 bp in length
5415 5514: gap of unknown length
5515 8773: contig of 3259 bp in length
8774 10904: contig of 2031 bp in length
10905 11004: gap of unknown length
11005 13204: contig of 2200 bp in length
13205 13304: gap of unknown length
13305 16401: contig of 3097 bp in length
16402 16501: gap of unknown length
16502 18227: contig of 1726 bp in length
18228 18327: gap of unknown length
18328 22401: contig of 4074 bp in length
22402 22501: gap of unknown length
22501 25177: contig of 2676 bp in length
25178 25277: gap of unknown length
25278 29892: contig of 4615 bp in length
29893 33457: gap of unknown length
33458 33557: gap of unknown length
33558 37405: contig of 3848 bp in length
37406 37505: gap of unknown length
37506 41734: contig of 4229 bp in length
41735 41834: gap of unknown length
41835 48989: contig of 7165 bp in length
48990 49099: gap of unknown length
49100 56076: contig of 6977 bp in length
56077 56176: gap of unknown length
56177 63604: contig of 7428 bp in length
63605 63704: gap of unknown length
63705 70821: contig of 7117 bp in length
70822 70921: gap of unknown length
70922 80138: contig of 9217 bp in length
80139 80238: gap of unknown length
80239 88780: contig of 8542 bp in length
88781 88880: gap of unknown length
88881 104642: contig of 15762 bp in length
104643 104742: gap of unknown length
104743 124656: contig of 19914 bp in length
124657 124756: gap of unknown length

* 124757 143891: contig of 19135 bp in length
* 143892 143991: gap of unknown length
* 143992 170586: contig of 26595 bp in length.
Location/Qualifiers

FEATURES

source

1. 170586

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="02"

/clone="RP11-422L5"

/clone_id="RP11-1"

1. 1171

/note="assembly_name:Contig1"

1272. 2319

/note="assembly_name:Contig13"

2420. 3793

/note="assembly_name:Contig17"

3894. 5414

/note="assembly_name:Contig18"

5515. 8773

/note="assembly_name:Contig19"

8874. 10904

/note="assembly_name:Contig20"

clone_end:17"

11005. 13204

/note="assembly_name:Contig21"

13305. 16401

/note="assembly_name:Contig22"

16502. 18227

/note="assembly_name:Contig23"

18328. 22401

/note="assembly_name:Contig24"

22502. 25177

/note="assembly_name:Contig25"

25278. 29892

/note="assembly_name:Contig26"

29993. 33457

/note="assembly_name:Contig27"

33558. 37405

/note="assembly_name:Contig28"

37506. 41734

/note="assembly_name:Contig29"

41835. 48999

/note="assembly_name:Contig30"

49100. 56076

/note="assembly_name:Contig31"

clone_end:SP6"

56177. 63604

/note="assembly_name:Contig32"

63705. 70821

/note="assembly_name:Contig33"

70922. 80138

/note="assembly_name:Contig34"

80239. 88780

/note="assembly_name:Contig35"

88881. 104642

/note="assembly_name:Contig36"

104743. 124656

/note="assembly_name:Contig37"

124757. 143891

/note="assembly_name:Contig38"

143992. 170586

/note="assembly_name:Contig39"

BASE COUNT 50417 a 34173 c 33290 g 50259 t 2447 others
ORIGIN

Query Match

26.7%; Score 414; DB 2; Length 170586;

Best Local Similarity 100.0%; Pred. No. 6.9e-194;

Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

187 AGTCCACGACAGTCGGTTGTAGATTCCTGCATCTGCAATTCCTCCGTTCCCTTCAT 246

Db

140127 AGTCCACGACAGTCGGTTGTAGATTCCTGCATCTGCAATTCCTCCGTTCCCTTCAT 140068

OY 247 GGATTGAAGCTCTCTCTTCTTCTTGAATGACTTTGCATCCCTCTGTTGTAGAG 306
 DB 140067 GGATTGAAGCTCTCTCTTCTTCTTGAATGACTTTGCATCCCTCTGTTGTAGAG 140008
 OY 307 TTGGACAAATGTTGATACGTACGTTGAGAACCAAGCCACACATACTGTAATACACTCTT 366
 DB 140007 TTGGACAAATGTTGATACGTACGTTGAGAACCAAGCCACACATACTGTAATACACTCTT 139948
 OY 367 CCTGACCAATGACCTGACTGAGAGAGATGATGAGAGAGGTGCTGCAAAAGAGGAGACT 426
 DB 139947 CCTGACCAATGACCTGACTGAGAGAGATGATGAGAGAGGTGCTGCAAAAGAGGAGACT 139888
 OY 427 CATCTCTCTTACCATCCGCTATCTTCCGACCATGAGAGGCATACCTGGACACATG 486
 DB 139887 CATCTCTCTTACCATCCGCTATCTTCCGACCATGAGAGGCATACCTGGACACATG 139828
 OY 487 GAAGAGAGGCTGCTGAGATCCGGCTGCTGAGAACAGAGTACTCTCTCTCATAC 546
 DB 139827 GAAGAGAGGCTGCTGAGATCCGGCTGCTGAGAACAGAGTACTCTCTCTCATAC 139768
 OY 547 AGCCATGATCTGCGGCCCAAGGGCGTCAACACTGTTGGCTAAAGGCTTGG 600
 DB 139767 AGCCATGATCTGCGGCCCAAGGGCGTCAACACTGTTGGCTAAAGGCTTGG 139714
 RESULT 12
 AC005037 190508 bp DNA linear PRI 30-SEP-2000
 LOCUS Homo sapiens BAC clone RP11-469M7 from 2, complete sequence.
 ACCESSION AC005037
 VERSION AC005037.2 GI:4827310
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE
 AUTHORS 2 (bases 1 to 190508)
 TITLE Abbot, A. and Le, T.
 JOURNAL The sequence of Homo sapiens BAC clone RP11-469M7
 REFERENCE
 AUTHORS 3 (bases 1 to 190508)
 TITLE Unpublished
 JOURNAL Waterston, R.H.
 REFERENCE
 AUTHORS Direct Submission
 TITLE Submitted (12-JUN-1998) Genome Sequencing Center, Washington
 JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE
 AUTHORS 4 (bases 1 to 190508)
 TITLE Waterston, R.H.
 JOURNAL Direct Submission
 TITLE Submitted (14-MAY-1999) Genome Sequencing Center, Washington
 JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE
 AUTHORS 5 (bases 1 to 190508)
 TITLE Waterston, R.
 JOURNAL Direct Submission
 TITLE Submitted (30-SEP-2000) Department of Genetics, Washington
 JOURNAL University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On May 14, 1999 this sequence version replaced g1:3309089.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: saplens@wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0469M07

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-1308; the clone sequenced to the right is RP11-91M5. Actual start of this clone is at base position 1 of RP11-469M7; actual end is at 190508 of RP11-469M7.

FEATURES

source	FEATURES
1..190508	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone_1id="RP11-469M7"
	/clone_1lb="RP11-11"
155..222	/rpt_family="12"
982..1299	/rpt_family="Alu"
1502..1533	/rpt_family="CA)n"
1535..1823	/rpt_family="Alu"
2005..2308	/rpt_family="Alu"
2392..2562	/rpt_family="Alu"
3284..3315	/rpt_family="L1"
3316..3338	/rpt_family="Alu"
3339..3621	/rpt_family="(CAA)n"
3650..3743	/rpt_family="Alu"
4353..4664	/rpt_family="AT-rich"
4727..5018	/rpt_family="Alu"
5545..5585	/rpt_family="Alu"
5586..5873	/rpt_family="(CAT)n"
5876..6007	/rpt_family="Alu"

repeat_region	repeat_family="MER21_g"	repeat_region	repeat_family="Alu"
repeat_region	6327. .6713	repeat_region	/rpt_family="Alu"
repeat_region	6924. .7294	repeat_region	/rpt_family="MER21_g"
repeat_region	8027. .8049	repeat_region	/rpt_family="MER21_type"
repeat_region	8050. .8354	repeat_region	/rpt_family="AT_r1ch"
repeat_region	8355. .8661	repeat_region	/rpt_family="Alu"
repeat_region	8356. .8684	repeat_region	/rpt_family="Alu"
repeat_region	9450. .9752	repeat_region	/rpt_family="AT_r1ch"
repeat_region	10371. .10680	repeat_region	/rpt_family="Alu"
repeat_region	10762. .11038	repeat_region	/rpt_family="Alu"
repeat_region	11039. .11109	repeat_region	/rpt_family="Alu"
repeat_region	11138. .11410	repeat_region	/rpt_family="Alu"
repeat_region	11446. .11663	repeat_region	/rpt_family="Alu"
repeat_region	11664. .11959	repeat_region	/rpt_family="L1"
repeat_region	11960. .12189	repeat_region	/rpt_family="Alu"
repeat_region	12190. .12334	repeat_region	/rpt_family="L1"
repeat_region	12335. .12360	repeat_region	/rpt_family="Alu"
repeat_region	12464. .12514	repeat_region	/rpt_family="L1"
repeat_region	12515. .12822	repeat_region	/rpt_family="L1"
repeat_region	12823. .12862	repeat_region	/rpt_family="Alu"
repeat_region	12857. .13096	repeat_region	/rpt_family="L1"
repeat_region	13101. .13330	repeat_region	/rpt_family="L1"
repeat_region	13331. .13634	repeat_region	/rpt_family="MER21_g"
repeat_region	13635. .13692	repeat_region	/rpt_family="Alu"
repeat_region	13704. .13990	repeat_region	/rpt_family="MER21_g"
repeat_region	14082. .14387	repeat_region	/rpt_family="L1"
repeat_region	14402. .14704	repeat_region	/rpt_family="Alu"
repeat_region	14797. .15095	repeat_region	/rpt_family="Alu"
repeat_region	15162. .15196	repeat_region	/rpt_family="Alu"
repeat_region	15606. .15681	repeat_region	/rpt_family="Mariner"
repeat_region	15691. .15827	repeat_region	/rpt_family="L1"
repeat_region	15861. .16189	repeat_region	/rpt_family="MER21_type"
repeat_region	16426. .16446	repeat_region	/rpt_family="L2"
repeat_region	16447. .16730	repeat_region	/rpt_family="L2"
repeat_region	17072. .17179	repeat_region	/rpt_family="Alu"
repeat_region	17212. .17504	repeat_region	/rpt_family="L2"
repeat_region	17869. .18139	repeat_region	/rpt_family="Alu"
repeat_region	18168. .18464	repeat_region	/rpt_family="Alu"

Query Match 26.7%; Score 414; DB 9; Length 190508; Best Local Similarity 100.0%; Pred. No. 6.7e-194; Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

repeat_region	repeat_family="Alu"
repeat_region	18606. .19155
repeat_region	19454. .19767
repeat_region	19807. .19949
repeat_region	19950. .20247
repeat_region	/rpt_family="MER21_type"
repeat_region	/rpt_family="Alu"

RESULT 13 AC093681 198250 bp DNA linear HTG 09-OCT-2001

LOCUS Homo sapiens chromosome 4 clone RP11-663N2, *** SEQUENCING IN PROGRESS ***; 26 unordered pieces.

ACCESSION AC093681.2 GI:15987253

VERSION AC093681

KEYWORDS HTG: HTGS_PHASE1

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 198250) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Waterston, R.H.

TITLE The sequence of Homo sapiens clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 198250) Waterston, R.H.

AUTHORS Direct Submission

TITLE Submitted (07-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

JOURNAL On Oct 9, 2001 this sequence version replaced gi:15487524.

COMMENT ----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@wustl.edu

Project Information -----

[illegible]

```
Center project name: H_NH063N02
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1912: contig of 1912 bp in length
* 1913 2012: gap of unknown length
* 2013 3539: contig of 1527 bp in length
* 3540 3639: gap of unknown length
* 3640 6127: contig of 2488 bp in length
* 6128 6227: gap of unknown length
* 6228 9244: contig of 3017 bp in length
* 9245 9344: gap of unknown length
* 9345 13288: contig of 3944 bp in length
* 13289 13388: gap of unknown length
* 13389 16301: contig of 2913 bp in length
* 16302 16401: gap of unknown length
* 16402 20096: contig of 3695 bp in length
* 20097 20196: gap of unknown length
* 20197 24467: contig of 4271 bp in length
* 24468 24567: gap of unknown length
* 24568 28680: contig of 4113 bp in length
* 28681 32494: contig of 3714 bp in length
* 32495 32594: gap of unknown length
* 32595 3526: contig of 3032 bp in length
* 35267 40050: contig of 4324 bp in length
* 40051 40150: gap of unknown length
* 40151 43881: contig of 3731 bp in length
* 43882 43981: gap of unknown length
* 43982 50194: contig of 6213 bp in length
* 50195 50294: gap of unknown length
* 50295 57454: contig of 7160 bp in length
* 57455 57554: gap of unknown length
* 57555 64660: contig of 7106 bp in length
* 64661 64760: gap of unknown length
* 64761 72159: contig of 7399 bp in length
* 72160 72259: gap of unknown length
* 72260 79835: contig of 7576 bp in length
* 79836 79935: gap of unknown length
* 79936 90863: contig of 10928 bp in length
* 90864 90964: gap of unknown length
* 90965 99400: contig of 8437 bp in length
* 99401 99500: gap of unknown length
* 99501 114190: contig of 14690 bp in length
* 114191 114290: gap of unknown length
* 114291 127822: contig of 13532 bp in length
* 127823 127922: gap of unknown length
* 127923 143474: contig of 15552 bp in length
* 143475 143574: gap of unknown length
* 143575 170444: contig of 26870 bp in length
* 170445 170544: gap of unknown length
* 170545 195651: contig of 25107 bp in length
* 195652 195751: gap of unknown length
* 195752 198250: contig of 2499 bp in length.
*
FEATURES
Source
1. 198250
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-663N2"
1. 1912
/note="assembly_name:Contig10"
misc_feature
2013..3539
/note="assembly_name:Contig11"
misc_feature
3640..6127
misc_feature
6228..9244
/note="assembly_name:Contig12"
misc_feature
9345..13288
/note="assembly_name:Contig13"
misc_feature
13389..16301
/note="assembly_name:Contig14"
misc_feature
16302..20096
/note="assembly_name:Contig15"
misc_feature
20197..24467
/note="assembly_name:Contig16"
misc_feature
24568..28680
/note="assembly_name:Contig17"
misc_feature
28781..32494
/note="assembly_name:Contig18"
misc_feature
32595..3526
/note="assembly_name:Contig19"
misc_feature
35267..40050
/note="assembly_name:Contig20"
misc_feature
40151..43881
/note="assembly_name:Contig21"
misc_feature
43982..50194
/note="assembly_name:Contig22"
misc_feature
50295..57454
/note="assembly_name:Contig23"
misc_feature
57555..64660
/note="assembly_name:Contig24"
misc_feature
64761..72159
/note="assembly_name:Contig25"
misc_feature
72260..79835
/note="assembly_name:Contig26"
misc_feature
79936..90863
/note="assembly_name:Contig27"
misc_feature
90964..99400
/note="assembly_name:Contig28"
misc_feature
99501..114190
/note="assembly_name:Contig29"
misc_feature
114291..127822
/note="assembly_name:Contig30"
misc_feature
127923..143474
/note="assembly_name:Contig31"
misc_feature
143575..170444
/note="assembly_name:Contig32"
misc_feature
170545..195651
/note="assembly_name:Contig33"
misc_feature
195752..198250
/note="assembly_name:Contig34"
misc_feature
56011 a 42355 c 41444 g 55928 t 2512 others
BASE COUNT
ORIGIN
Query Match 26.7% Score 414; DB 2; Length 198250;
Best Local Similarity 100.0%; Pred. No. 6,78-194;
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 AGTCCCGACGACGACGCGGTTGTGATTCCTGATCTGCAATTTCCCGTCCCTTAT 246
|||||
DB 14089 AGTCCCGACGACGACGCGGTTGTGATTCCTGATTCCTGCAATTTCCCGTCCCTTAT 14148
QY 247 GGAATTGAAGGCTCCTCTTCTCTTGAATGACCTTGCAATTCCTGCTGAGAG 306
|||||
DB 14149 GGAATTGAAGGCTCCTCTTCTCTTGAATGACCTTGCAATTCCTGCTGAGAG 14208
QY 307 TTGGGACATGTTGGATTACTGTTGGAACCAAGCCACACATCTGTAATATACACTTT 366
|||||
DB 14209 TTGGGACATGTTGGATTACTGTTGGAACCAAGCCACACATCTGTAATATACACTTT 14268
QY 367 CCTGACCAATGACTGACTGAGGAAGTGTGAGAGAGTGCTGCAAAAAGAGGAGACT 426
|||||
DB 14269 CCTGACCAATGACTGACTGAGGAAGTGTGAGAGAGTGCTGCAAAAAGAGGAGACT 14328
QY 427 CATTCCTGCTACATTCGCGCTATCTCGACCCATGAGGCGCATTAACCTGGACACATG 486
|||||
DB 14329 CATTCCTGCTACATTCGCGCTATCTCGACCCATGAGGCGCATTAACCTGGACACATG 14388
```

QY 487 GAAGAGGCGCTGGTGTATCCGGGCTGTGGAGAACAGTGGTATCTACTCTCCCTATAC 546
DB 14389 GAAGAGGCGCTGGTGTATCCGGGCTGTGGAGAACAGTGGTATCTACTCTCCCTATAC 14448
QY 547 AGCCTATGATGCTGGCGCCAGGCGCTCAACAACACTGGTGGCTAAAGGCGCTGG 600
DB 14449 AGCCTATGATGCTGGCGCCAGGCGCTCAACAACACTGGTGGCTAAAGGCGCTGG 14502

RESULT 14
BD049005 249 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD049005
VERSION BD049005.1 GI:22590747
KEYWORDS JP 2001269182-A/25251.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 249)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 25251 02-OCT-2001;
GENSET

COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/25251
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PI JORDAN
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC

FEATURES
source 1..249 location/Qualifiers.
FH Key location/Qualifiers.
1..249
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 62 a 65 c 57 g 63 t 2 others
ORIGIN

Query Match 13.5%; Score 210; DB 6; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.6e-92;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 TTGAATGACCTTGCATCCCTCGTTCGTGAGAGTTGGACAATGTTGATTAAGTGTG 331
DB 37 TTGAATGACCTTGCATCCCTCGTTCGTGAGAGTTGGACAATGTTGATTAAGTGTG 96
QY 332 GAACCAAGCCACACATACCTGTAATACACTCTCTGACCAATGACCTGATGAGAA 391
DB 97 GAACCAAGCCACACATACCTGTAATACACTCTCTGACCAATGACCTGATGAGAA 156
QY 392 GTGATGAGAGAGTGTGCAAAAAGAGGAGACCTCATTTCTCTCAACATCCGCTATC 451
DB 157 GTGATGAGAGAGTGTGCAAAAAGAGGAGACCTCATTTCTCTCAACATCCGCTATC 216
QY 452 TTCCGACCATGAAAGGCAATACCTGGAC 481
DB 217 TTCCGACCATGAAAGGCAATACCTGGAC 246

RESULT 15
AC093681/c 198250 bp DNA linear HTG 09-OCT-2001
LOCUS Homo sapiens chromosome 4 clone RP11-663N2, *** SEQUENCING IN
DEFINITION PROGRESS ***, 26 unordered pieces.

ACCESSION AC093681
VERSION AC093681.2 GI:15987253
KEYWORDS HTG; HTGS_PHASE1
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 198250)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
2 (bases 1 to 198250)
REFERENCE Waterston,R.H.
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Oct 9, 2001 this sequence version replaced gi:15487524.

COMMENT ----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: H.NH0663N02

* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1912: contig of 1912 bp in length
* 1913 2012: gap of unknown length
* 2013 3539: contig of 1527 bp in length
* 3540 3639: gap of unknown length
* 6127: contig of 2488 bp in length
* 6128 6227: gap of unknown length
* 6228 9244: contig of 3017 bp in length
* 9245 9345: gap of unknown length
* 9345 13288: contig of 3944 bp in length
* 13289 13388: gap of unknown length
* 13389 16301: contig of 2913 bp in length
* 16302 16401: gap of unknown length
* 16402 20096: contig of 3695 bp in length
* 20097 20197: gap of unknown length
* 20197 24467: contig of 4271 bp in length
* 24468 24567: gap of unknown length
* 24568 28680: contig of 4113 bp in length
* 28681 28780: gap of unknown length
* 28781 32494: contig of 3714 bp in length
* 32495 32595: gap of unknown length
* 32595 35626: contig of 3032 bp in length
* 35627 35726: gap of unknown length
* 35727 40050: contig of 4324 bp in length
* 40051 40150: gap of unknown length
* 40151 43881: contig of 3731 bp in length
* 43882 43981: gap of unknown length
* 43982 50194: contig of 6213 bp in length
* 50195 50294: gap of unknown length
* 50295 57454: contig of 7160 bp in length
* 57455 57534: gap of unknown length
* 57535 64660: contig of 7106 bp in length
* 64661 72159: contig of 7399 bp in length
* 72160 72259: gap of unknown length
* 72260 79835: gap of 7576 bp in length
* 79836 79935: gap of unknown length
* 79936 90863: contig of 10928 bp in length
* 90864 gap of unknown length

* 90964 99400: contig of 8437 bp in length
* 99401 99500: gap of unknown length
* 99501 114190: contig of 14690 bp in length
* 114191 114290: gap of unknown length
* 114291 127822: contig of 13532 bp in length
* 127823 127922: gap of unknown length
* 127923 143474: contig of 15552 bp in length
* 143475 170444: gap of unknown length
* 170445 170544: contig of 26870 bp in length
* 170545 195651: gap of unknown length
* 195652 195751: gap of unknown length
* 195752 198250: contig of 2499 bp in length.
* Location/Qualifiers
1. 198250
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-663N2"
1. 1912

misc_feature /note="assembly_name:Contig10"
2013. 3539
misc_feature /note="assembly_name:Contig11"
3640. 6127
misc_feature /note="assembly_name:Contig12"
6228. 9244
misc_feature /note="assembly_name:Contig13"
9345. 13288
misc_feature /note="assembly_name:Contig14"
13389. 16301
misc_feature /note="assembly_name:Contig15"
16402. 20096
misc_feature /note="assembly_name:Contig16"
20197. 24467
misc_feature /note="assembly_name:Contig17"
24568. 28680
misc_feature /note="assembly_name:Contig18"
28781. 32494
misc_feature /note="assembly_name:Contig19"
32595. 35626
misc_feature /note="assembly_name:Contig20"
35727. 40050
misc_feature /note="assembly_name:Contig21"
40151. 43881
misc_feature /note="assembly_name:Contig22"
43982. 50194
misc_feature /note="assembly_name:Contig23"
50295. 57454
misc_feature /note="assembly_name:Contig24"
57555. 64660
misc_feature /note="assembly_name:Contig25"
64761. 72159
misc_feature /note="assembly_name:Contig26"
72260. 79835
misc_feature /note="assembly_name:Contig27"
79936. 90863
misc_feature /note="assembly_name:Contig28"
90964. 99400
misc_feature /note="assembly_name:Contig29"
99501. 114190
misc_feature /note="assembly_name:Contig30"
114291. 127822
misc_feature /note="assembly_name:Contig31"
127923. 143474
misc_feature /note="assembly_name:Contig32"
143575. 170444
misc_feature /note="assembly_name:Contig33"
170545. 195651
misc_feature /note="assembly_name:Contig34"
195752. 198250
misc_feature /note="assembly_name:Contig9"
56011 a 42355 c 41444 g 55928 t 2512 others

ORIGIN
Query Match 9.1%; Score 141; DB 2; Length 198250;
Best Local Similarity 100.0%; Pred. No. 1.2e-58;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 888 AGCTTTGCTTCTACATACATGGAATGGGAGCGTTATGCACACTGGATGAATCTGTCTCC 947
DB 10961 AGCTTTGCTTCTACATACATGGAATGGGAGCGTTATGCACACTGGATGAATCTGTCTCC 10902
QY 948 TGGCAACCATGATGATGCAATTAAGACACCTAAACCTATCTATTCGCTTAGCCC 1007
DB 10901 TGGCAACCATGATGATGCAATTAAGACACCTAAACCTATCTATTCGCTTAGCCC 10842
QY 1008 TTGGGGTGGGAGAACCTTAG 1028
DB 10841 TTGGGGTGGGAGAACCTTAG 10821
Search completed: August 23, 2003, 16:41:11
Job time : 3963 secs